

**Б. В. Орехов**

*Национальный исследовательский университет*

*«Высшая школа экономики»*

*(Россия, Москва)*

*nevmenandr@yandex.ru*

## **ЕЩЕ РАЗ ОБ ИССЛЕДОВАТЕЛЬСКОМ ПОТЕНЦИАЛЕ ПОЭТИЧЕСКОГО КОРПУСА: МЕТР, ЛЕКСИКА, ФОРМУЛА**

Статья продолжает начатую в публикациях других исследователей демонстрацию возможностей поэтического подкорпуса в составе Национального корпуса русского языка. Речь идёт о том, какие вопросы, связанные с историей русской поэзии, можно решать с помощью корпуса. В первой части статьи проводится пилотное исследование тяготения лексических единиц к определенным поэтическим метрам. Исследование проводится с применением коллострукционного анализа. В качестве тестового материала используются «метапоэтические термины», то есть такие, которые традиционно используются при описании процесса создания поэтического произведения. Почти все они оказываются тяготеющими к ямбу. Исключение составляет слово «хорей», появляющееся чаще в хорейских строках. Вторая часть статьи реконструирует поэтическую традицию, в которой находится стихотворение Ф.И. Тютчева «Последний катаклизм». Стихотворение рассматривается как составленное из ряда поэтических формул, повторяемых текстовых фрагментов, каждая из которых имеет свою историю, происхождение и частотность в русской поэзии. Уяснение этих аспектов для каждой формулы помогает полнее представить генезис всего стихотворения и встроить его в модель восприятия читателя-современника.

*Ключевые слова:* корпусная лингвистика, поэзия, русская поэзия, поэтический метр, поэтическая лексика.

Эта статья продолжает начатую в публикациях [Гришина и др. 2009], [Сичинава 2012] демонстрацию исследовательских возможностей поэтического подкорпуса в составе Национального корпуса русского языка. Если в упомянутых статьях в равной степени отмечаются способы использования поэтического подкорпуса для решения как чисто лингвистических задач (например, диахроническое описание акцентных вариантов), так и связанных с собственно историей русской поэзии, то мы сосредоточимся на втором круге вопросов, не претендуя при этом охватить его целиком.

Поэтический корпус предоставляет уникальную возможность изучения взаимодействий метра и различных лингвистических явлений (например, поведения лексических единиц) в стихе. Эта возможность появляется благодаря тому, что к привычным лингвистическим разметкам (морфологической, семантической, метатекстовой) добавлена специальная метрическая разметка, то есть каждой стихотворной строке приписана помета, определяющая метр, стопность и клаузулу стиха. Поискные механизмы корпуса позволяют задавать пользовательские подкорпуса, включающие тексты, состоящие из строк, отвечающих требуемым характеристикам.

## 1. Метр и лексика

Получив такой инструмент, исследователь оказывается способен поставить вопрос о наличии неочевидных закономерностей взаимного тяготения лексики и метра. М. Л. Гаспаров оформил в виде последовательной теории «семантического ореола метра» высказывавшиеся ранее специалистами по русскому стиху идеи о системной связи между темой (и даже набором конкретных образов) стихотворения и избираемым для текста метром [Гаспаров 2013]. Так как тема текста в конечном счете формируется конкретными словами, гипотеза о том, что слово и метр могут формировать исторически обусловленную связь, выглядит логичной, а корпус предоставляет возможности для ее проверки.

М. Л. Гаспаров подчеркивал, что наличие семантического ореола не означает строгой последовательности в его реализации в тексте. Иными словами, классический пример тематического тяготения — тема дороги/пути в стихотворениях, написанных пятистопным хореем, может отсутствовать и, напротив, обнаруживать себя в сти-

хотворениях, написанных другим метром. Таким образом, исследователю приходится иметь дело с нечеткими границами и критериями анализируемого явления.

Однако такое нечеткое на первый взгляд понятие как «тяготение» вполне может быть осмыслено в терминах современной статистики. Задачи, связанные с нахождением взаимного влияния ряда факторов, скажем, в квантитативной лингвистике, решаются с помощью коллострукционного анализа, являющегося расширением традиционного критерия Фишера (подробнее см. Gries, Stefanowitsch 2004; Stefanowitsch, Gries 2003; Кузнецова, Велейшикова 2010)<sup>1</sup>. Эта статистическая методика позволяет устанавливать наличие «притяжения», «отталкивания» или индифферентности двух явлений (например, слов) в случае их совместной встречаемости в выборке.

Собрав материал в поэтическом подкорпусе, мы можем построить матрицу частотности лексем в строках определенного метра и вычислить зависимость или независимость слова от метра в русской поэзии. Ручная работа такого рода была бы чересчур объемной и ее выполнимость за разумное время без корпуса можно поставить под сомнение.

Для демонстрации метода мы избрали двадцать лексем, которые условно можно отнести к метапоэтическим терминам, то есть к классу слов, так или иначе являющихся метонимическими заместителями процесса поэтического творчества: *поэт, лира, поэзия, стих, муза, песнь, вдохновение, перо, слагать, сонет, пиит, элегия, ода, ямб, хорей, гомер, алкей/алцей, тетрадь, строка, строчка*.

Приведем ряд контекстов с этими лексемами в поэтическом корпусе:

Поэт-мечта дремотного Гиганта,  
Спит Исполин и слышит звоны **лир**.

[К. Д. Бальмонт. «Задуман безошибочным Творцом...» [Два сонета к Италии, 2] (1892–1935)]

Гордяся им, я посвящаю вам  
Игривый звук заздравного **сонета**,

<sup>1</sup> Автор благодарит Ю. Л. Кузнецову за консультации в этом вопросе.

Моей души усердный фимиам.

[Н. М. Языков. «На праздник ваш принес я два привета...»  
(1829–1833)]

А рядом здесь, под миртовой кушей,

Еще звучит, полна любовной неги,

Гармония Тибулловых **элегий**...

[Д. С. Мережковский. Тибур (1891)]

Жизнь сердца для него единственный предмет

Всех песней пламенных всех томных **вдохновений**;

Жизнь сердца кончится, — в молчаньи и смиреньи

Он укрывается... о, да!

[Е. П. Ростопчина. И он поэт (1840.05.09)]

Интерпретация лексемы *тетрадь* как метапоэтического термина может быть воспринята с сомнением современным читателем, однако контексты, извлеченные из поэтического корпуса, надежно подсказывают, что в стихотворном тексте слово *тетрадь* употребляется прежде всего как место для фиксации творческой (прежде всего, поэтической) активности человека.

Нам вчера

В редакцию товарищ прокурора

Прислал **тетрадь** рифмованного здра.

[Д. С. Мережковский. Вера (1890)]

Я медленно закрыл **тетрадь**,

Перо подальше отложил

И начал терпеливо ждать,

Чтоб мрачный гость заговорил...

[С. Я. Надсон. Песни Мефистофеля (1886)]

А подле — с пестрыми каймами

*Тетради* песен и стихов

Счастливой Персии певцов...

[И. И. Козлов. Абидосская невеста (1826)]

Так как нас интересует плоскость семантики стиха, мы попытаемся провести анализ функционирования в поэтическом тексте именно лексем, а не конкретных словоформ. Таким образом, за пределами нашего исследования остаются все просодические особенности слов, проявляющие себя при словоизменении. Так, наращение слогов (*сонет — сонета*), сокращение длины слова (*муза — муз*) или смещение ударения (*стих — стихов*) при склонении в нашей выборке нивелируются. Это особенно важно осознавать, так как интуитивно может показаться, что слова ямбической структуры (с ударением на второй слог) должны тяготеть к ямбу, а слова хорейской структуры (с ударением на первый слог) должны тяготеть к хорю. Однако это было бы заблуждением: граница полноударных слов в русском стихе не обязательно совпадает с границами стопы. В следующем примере двусложная словоформа *поэт* занимает только часть трёхсложной стопы анапеста:

Но, деляся с толпой этой тайной души сокровенной,  
Изменяет себе, изменяет ли тайне **поэт**...

[О. Н. Чюмина. «Говорят, что порой, совлекая бесстрашно покровы...» (1893)]

В следующем примере словоформа *поэт* оказывается на границе двух дактилических стоп:

Я — не **поэт**, но с любовью глубокой  
Музе-шалунье, как эхо, служу...

[Л. Н. Трефолев. Пляска весны (1892.01.20)]

Извлеченные из корпуса данные не нормализованы по авторам, так что остаются основания предполагать, что обнаруженные закономерности обусловлены не традицией русской поэзии, а индивидуальными стилевыми предпочтениями авторов, которые способны влиять на распределение. В то же время корпус объемом в 9,6 млн. словоупотреблений позволяет надеяться на сбалансированность исследованного материала и сглаженность влияния случайных факторов.

В ходе исследования нами последовательно были заданы подкорпуса для каждого из пяти метров русской силлабо-тоники. В каж-

дом из подкорпусов мы провели поиск интересующей нас лексемы. Результаты предварительных замеров приведены в Таблице 1.

Лексема	Я	Х	Д	Ам	Ан	Всего
поэт	3413	575	99	246	181	5603
лира	1440	255	37	66	37	2078
поэзия	707	62	19	40	27	1027
стих	4396	621	84	202	161	6758
муза	2048	243	70	56	55	2797
песнь	3051	1196	325	385	339	6577
вдохновение	210	28	4	9	8	294
перо	965	265	22	57	71	1808
слагать	247	35	10	24	12	381
сонет	487	28	2	14	9	565
пиит	242	13	10	1	7	292
элегия	289	26	4	10	7	352
ода	552	100	10	4	15	758
ямб	88	10	1	11	4	144
хорей	14	20	0	3	2	50
гомер	147	17	4	4	5	238
алкей/алцей	5	0	0	0	0	10
тетрадь	197	31	8	10	14	313
строка	617	118	24	56	44	1131
строчка	132	57	7	27	17	388
подкорпус	5205366	1386732	225491	415453	402174	9671137

Таблица 1. Частотность лексем в подкорпусах метров

Особенность полученных данных в бросающемся в глаза превалировании ямбического подкорпуса над остальными подкорпусами метров. Действительно, ямб — самый распространенный метр русской поэзии. Из-за разницы в объемах корпусов, подчас отличающихся по числу словоупотреблений на порядок, определить тяготение «на глаз» было бы крайне трудно.

Результаты подсчетов с использованием методики коллострукционного анализа приведены в Таблице 2. Плюсом обозначено тя-

готение слова к метру, минусом отталкивание, «0» маркирует нейтральность.

Лексема	Я	Х	Д	Ам	Ан
поэт	+	-	-	0	-
лира	+	-	0	-	-
поэзия	+	-	0	0	-
стих	+	-	-	-	-
муза	+	-	0	-	-
песнь	-	+	+	+	+
вдохновение	+	-	0	0	0
перо	0	0	-	-	0
слагать	+	-	0	0	0
сонет	+	-	-	-	-
пиит	+	-	0	-	0
элегия	+	-	0	0	-
ода	+	0	0	-	-
ямб	0	-	0	0	0
хорей	-	+	0	0	0
гомер	+	-	0	0	0
алкей/алцей	0	0	0	0	0
тетрадь	+	-	0	0	0
строка	0	-	0	0	0
строчка	-	0	0	+	0

Таблица 2. Притяжения, отталкивания и нейтральность лексем и метров по данным коллострукционного анализа

Действительно, ряд метапоэтических терминов обнаружил свое тяготение к определенным метрам. Более всего попавшие в выборку лексемы тяготеют к ямбу: 13 из 20 слов, скорее, следует ожидать увидеть в ямбических строках. При этом статистический курьез в том, что сама лексема *ямб* к ямбу как метру нейтральна, «отталкивается» от хореических стихов. Слово *хорей*, напротив, несмотря на ямбическую просодическую структуру, тяготеет к хореическим строкам и с наименьшей вероятностью встретится в ямбах.

Розы крови, рея,  
Дали красный нимб...  
Ровный бег **хорея**  
Подарил Олимп.  
[Е. И. Дмитриева. Хореи (1915)]

Любопытно и обнаруженное тяготение лексемы *строчка* к амфибрахию (независимо от стопности):

Альбом мой наполнен: всё врезано, вклеено;  
Хотела еще бы хоть **строчку** от Мея, но...  
[Л. А. Мей. (В альбом жене) (1858)]

Достойная **строчек**  
О мраморе рук,  
Вы платя комочек  
Вздымаете вдруг,  
[А. П. Ладинский. Как север (1930)]

Не сразу, не всем и не громко  
должна я об этом сказать,  
то **строчкой** мальчишески ломкой  
то просто поглядом в глаза.  
[О. Ф. Берггольц. «Мне многое в мире открыто...» (1928)]

Отдельного разбирательства требует случай слова *песнь*, отталкиваемого от ямба, но притягиваемого всеми остальными размерами. Можно выдвинуть еще нуждающуюся в уточнениях гипотезу, что хорей и разнообразные трехсложные размеры в большей степени закреплены за песенным дискурсом, чем «книжный» ямб (мы сознаем, что это утверждение найдет достаточное количество контрпримеров) и это могло повлиять на распределение лексемы в соответствующих подкорпусах.

## 2. Традиция и формулы

Одной из важнейших составляющих научной стратегии в гуманитарной сфере является установление генезиса описываемого яв-



ления в аспекте его взаимодействия с традицией. В филологии эта стратегия реализует себя в поиске «интертекстуальных параллелей» или источников культурно значимого текста. Поэтический корпус в этом смысле предоставляет наилучший инструмент для такого рода изысканий (см., в частности, Орехов 2013, Орехов 2014). Возможно, в будущем в корпусе может появиться разметка, автоматически маркирующая воспроизводимые фрагменты текста.

Мы называем «параллели» формулами (в духе Кавелти 1996), чтобы не делать различий между сознательным интертекстом и фразеологизованными ритмико-синтаксическими клише, так как в науке о литературе нет надежных способов установления «сознательности» использования автором того или иного приема.

Так как поэтический корпус содержит большинство значимых стихотворных текстов со второй трети XVIII века по середину XX века, поисковые механизмы позволяют легко определять степень распространенности той или иной конструкции, а также ее происхождение и ассоциированность с тем или иным авторским стилем.

Для примера можно, оттолкнувшись от стихотворения Ф. И. Тютчева «Последний катаклизм», проследить какая часть этого небольшого текста является формульной, а какие сочетания слов в контексте русской поэзии представляются оригинальными. При поиске в корпусе мы ограничили время создания попавших в выдачу текстов датой создания «Последнего катаклизма», 1829-м годом.

#### ПОСЛЕДНИЙ КАТАКЛИЗМ

Когда пробьет последний час природы,  
Состав частей разрушится земных:  
Всё зримое опять покроют воды,  
И божий лик изобразится в них!  
<1829>

Слово «катаклизм» до Тютчева в стихах не встречается. Более того, не упоминается оно и в текстах основного корпуса примерно до середины XIX века, по всей видимости, выступая в качестве экзотического «ученого» слова, привлекаемого поэтами барочного мироощущения (см. об этом Шайтанов 1998). Зато формула «когда пробьет» в инициальной позиции стиха не оригинальна и встречается до Тютчева, как минимум, однажды:

**Когда пробьет** печальный час  
Полночной тишины  
И звезды трепетно горят,  
Туман кругом луны, —  
[И. И. Козлов. Ирландская мелодия (1824)]

Тебя умерший не вспомнит,  
**Когда пробьет** кончины час;  
Безмолвен ад, — во тьме не станет  
Звучать Тебе хвалебный глас.  
[Д. П. Ознобишин. Псалом 6 (1820–1840)]

Сочетание «последний час» весьма частотно: до Тютчева оно использовалось не менее 44 раз. А вот конструкция «последний N природы» встречается в корпусе до Тютчева однажды в тексте автора, по признанию литературоведов, сильно повлиявшего на его творчество, что заставляет подозревать в сочетании «последний час природы» сознательное заимствование (однако доказать это, как нам представляется, невозможно):

Представь **последний** день природы,  
Что пролилася звезд река...  
[Г. Р. Державин. На взятие Измаила (1790–1791)]

Сочетание «состав частей» до Тютчева в русской поэзии не фигурирует (разумеется, по данным корпуса), а «части земные» употреблены в оде В. Майкова, однако, в другом значении (у Майкова 'части света', у Тютчева — 'составляющие материального мира'):

Посреди морской пучины  
И во всех **частях земных**,  
Где сражаются срацины...  
[В. И. Майков. Ода на день брачного сочетания цесаревича великого князя Павла Петровича и великия княгини Наталии Алексеевны (1773)]

Конструкция «все зримое» до Тютчева в корпусе не встречается, а сочетание лексем «покрыть» и «вода» в одном контексте сравни-

тельно частотно. Особенное пристрастие к нему имел также повлиявший на стиль Тютчева М. В. Ломоносов: из 12 удовлетворявших запросу контекстов, которые нашлись в корпусе, 5 принадлежат этому автору.

И что надуты **вод** громады  
Текли **покрыть** пространны грады,  
Сравнять хребты гор с влажным дном.

[М. В. Ломоносов. Ода на день восшествия на всероссийский престол ее величества государыни императрицы Елисаветы Петровны самодержицы всероссийския 1746 года (1746)]

**Покрыв** водами высоты,  
На легких облаках восходишь,  
Крилами ветров шум наводишь,  
Когда на них летаешь ты.

[М. В. Ломоносов. Преложение псалма 103 (1743–1749)]

Сочетание определяемого «лик» с определением «божий» в поэтическом корпусе до Тютчева встречается лишь однажды с инвертированным порядком в тексте М. М. Хераскова:

На лики божии взираю беспристрастно,  
Что в чем-нибудь грешна, сие мне кажет ясно.

[М. М. Херасков. Венецианская монахиня (1758)]

Глагол «изобразиться» в корпусе сравнительно частотный (16 вхождений), в одном контексте с водами, в которых и происходит «изображение» он встречается у того же М. М. Хераскова:

Какою горестью душа твоя разилась,  
Когда судьба твоя тебе **изобразилась**,  
Когда взглянула ты ко берегу шумных вод,  
Где вокруг твоих судов стесняется народ!

[М. М. Херасков. Россиада. Поэма эпическая. (1771–1779)]

Но в целом это, по всей видимости, устаревающий глагол, характерный, в основном, для поэтического языка XVIII века. После

Тютчева можно наблюдать еще 4 вхождения лексемы *изобразиться* в корпусе и ее полное исчезновение для поэзии XX века.

Таким образом, по контрасту с «Проблеском», в языковом смысле настроенном на сентиментальную поэтику В. А. Жуковского (подробнее см. об этом Орехов 2014), «Последний катаклизм» ориентирован на язык одической поэзии XVIII века, что выражается в избираемых Тютчевым формулах, призванных вызвать у читателя-современника соответствующий круг культурных ассоциаций.

Разумеется, исследовательский потенциал поэтического корпуса не ограничивается названными направлениями. Корпус представляет собой научный инструмент, идеально реализующий идею «поиск как язык», которая отражает два важнейших положения:

1. На языке можно порождать неопределённо большое множество заранее не заданных высказываний;

2. Создатель языка заранее не знает, что на нём будет выражено.

Аналогичным образом и поисковые инструменты в идеале должны быть достаточно гибкими и разнообразными, чтобы не ограничивать пользователя заранее придуманными создателем ресурса исследовательскими вопросами. Поэтический корпус полностью соответствует этому требованию и может быть использован в разнообразных лингвистических, стиховедческих и историко-культурных разысканиях.

## Литература

*Гаспаров М. Л.* Метр и смысл. Об одном из механизмов культурной памяти. М.: Фортунa ЭЛ, 2013.

*Гришина Е. А., Корчагин К. М., Плунгян В. А., Сичинава Д. В.* Поэтический корпус в рамках НКРЯ: общая структура и перспективы использования // Национальный корпус русского языка: 2006–2008. Новые результаты и перспективы. СПб.: Нестор-История, 2009, 71–113.

*Кавелти Дж.* Изучение литературных формул // Новое литературное обозрение. 1996. № 22. С. 33–64.

*Кузнецова Ю. Л., Велейшикова Т. В.* Современные корпусные исследования языка: новые подходы // Вопросы языкознания. 2010. № 6. С. 108–124.

Орехов Б. В. «Не Аполлон, но я кую сей цепи звенья...»: поздние стихи Батюшкова в свете корпусных данных // Корпусный анализ русского стиха. М.: Издательский центр «Азбуковник», 2013. С. 157–171.

Орехов Б. В. «Проблеск» Ф. И. Тютчева в ретроспективе Корпуса. Очерк корпусной поэтики // Корпусный анализ русского стиха: Сборник научных статей. Вып. 2. М.: Издательский центр «Азбуковник», 2014. С. 304–318.

Сичинава Д. В. Поэтический подкорпус Национального корпуса русского языка: несколько примеров поиска стиховедческой информации // Славянский стих. М.: Рукописные памятники Древней Руси, 2012. С. 482–491.

Шайтанов И. О. Уравнение с двумя неизвестными (Поэты-метафизики Джон Донн и Иосиф Бродский) // Вопросы литературы. 1998. № 6. С. 3–40.

Stefanowitsch A., Gries S. Th. Co-varying collexemes in the into-causative // Language, culture, and mind. Stanford, 2004. P. 225–236

Stefanowitsch A., Gries S. Th. Collostructions: investigating the interaction between words and constructions // International journal of corpus linguistics. 2003. 8 (2). P. 209–243

**B. V. Orekhov**

*National Research University Higher School of Economics*

*(Russia, Moscow)*

*nevmenandr@yandex.ru*

## **ONCE AGAIN ABOUT THE RESEARCH POTENTIAL OF THE POETIC CORPUS: METER, VOCABULARY, FORMULA**

The article continues the trend of other researchers' publications that demonstrate the opportunities of the poetic subcorpus of the Russian National corpus. The question is, what issues related to the history of Russian poetry can be solved with the help of the corpus. In the first part of the article there is a pilot study of attraction of lexical units to poetic meters. The study is conducted with the use of collostructional analysis. As the test material we use “metapoetics terms” — those which are traditionally used when describing the process of creating poetry.

Almost all of them are attracted to iambic verse. The only exception is the word “trochee” that appears more frequently in trochaic lines. The second part of the article reconstructs the poetic tradition, which a poem by Fyodor Tyutchev “Last cataclysm” follows. The poem is considered as composed of a set of poetic formulas (repeated text fragments), each of them with its own history, origin and frequency in Russian poetry. Clarifying these aspects for each formula helps to represent the genesis of the whole poem and embed it in the model of perception of the contemporary reader.

*Key words:* corpus linguistics, poetry, Russian poets, poetic meter, poetic vocabulary

### References

Cawelti J. [The study of literary formulas] *Novoe literaturnoe obozrenie* [New literary review]. Issue 22. Moscow, 1996. pp. 33–64. (In Russ.)

Gasparov M. L. *Metri i smysl. Ob odnom iz mehanizmov kul'turnoi pam'ati* [Meter and Meaning. One of the mechanisms of cultural memory]. Moscow, Fortuna EL Publ., 2013.

Grishina E. A., Korchagin K. M., Plungyan V. A., Sichinava D. V. [Poetic corpus in the RNC: General structure and prospects of use]. *Natsional'nyi korpus russkogo yazyka: 2006–2008. Novye rezul'taty i perspektivy* [Russian National Corpus: 2006–2008. New Results and Prospects]. St. Petersburg, Nestor-Istoriya Publ., 2009, pp. 71–113.

Kuznetsova Yu. L., Veleynshnikova T. V. [Modern corpus-based language studies: new approaches]. *Voprosy Yazykoznaniiya*, 2010, Issue 6, Moscow, pp. 108–124. (In Russ.)

Orekhov B. V. [“Not Apollo, but I make this chain links...”: late Batiushkov poems sub specie of corpus data]. *Korpusnyi analiz russkogo stikha* [Russian Verse corpus analysis]. Moscow, Azbukovnik Publ, 2013. pp. 157–171. (In Russ.)

Orekhov B. V. [“Glimpse” of F. I. Tyutchev in Corpus Retrospect. Essay on the Corpus poetics] *Korpusnyi analiz russkogo stikha* [Russian Verse corpus analysis]. Issue 2. Moscow, Azbukovnik Publ, 2014. pp. 304–318. (In Russ.)

Sichinava D. V. [Poetic subcorpus of the National corpus of the Russian language: a few examples of verse study relevant information].

*Slavjanskii stikh* [Slavic verse]. Moscow, Old Russian Manuscripts Publ., 2012, pp. 482–491. (In Russ.)

Shaytanov I. O. [Equation with two unknowns (the metaphysical Poets John Donne and Joseph Brodsky)]. *Voprosy literatury* [Questions of Literature]. Issue 6. Moscow, 1998. pp. 3–40. (In Russ.)

Stefanowitsch, A., Gries, S. Th. Co-varying collexemes in the into-causative. *Language, culture, and mind. Stanford*, 2004, pp. 225–236

Stefanowitsch, A., Gries, S. Th. Collostructions: investigating the interaction between words and constructions. *International journal of corpus linguistics*. 2003. 8 (2), pp. 209–243